

Kalman Filtering With Relays Over Wireless Fading Channels

Alex S. Leong and Daniel E. Quevedo

Abstract—This note studies the use of relays to improve the performance of Kalman filtering over packet dropping links. Packet reception probabilities are governed by time-varying fading channel gains, and the sensor and relay transmit powers. We consider situations with multiple sensors and relays, where each relay can either forward one of the sensors' measurements to the gateway/fusion center, or perform a simple linear network coding operation on some of the sensor measurements. Using an expected error covariance performance measure, we consider optimal and suboptimal methods for finding the best relay configuration, and power control problems for optimizing the Kalman filter performance. Our methods show that significant performance gains can be obtained through the use of relays, network coding and power control, with at least 30-40% less power consumption for a given expected error covariance specification.

I. INTRODUCTION

Since the seminal work of [2], which showed the existence of a critical threshold on the packet arrival rate for stability of the Kalman filter with i.i.d. Bernoulli packet losses, the problem of Kalman filtering over unreliable channels has received considerable attention. Extensions of [2] include the case of multiple sensors [3], [4], further characterizations of the critical threshold [5], probabilistic notions of performance [6], the consideration of delays [7], correlated packet losses [4], [8], and power control for stability [9].

In digital communications, channel coding is often used to improve the quality of transmissions over unreliable channels. The concept of network coding [10], [11], where in a network with many nodes, throughput can be increased by allowing intermediate nodes to perform simple operations (such as linear transformations [10]) on its received information, has attracted much attention in recent years. Kalman filtering with power control and coding was considered in [12], [13]. The recent work [13] included a study of network coding, where one could choose to utilize a relay to perform a network coding operation, and the energy tradeoffs involved. The use of relays in combating the effects of fading and increasing channel capacity has been extensively studied in wireless communications, see e.g. [14], [15]. Indeed, cooperative communications via the use of relays has been identified as one of the key enabling technologies for fifth generation (5G) mobile networks [16]. The use of a relay in control has been studied in [17], which showed that for the case of a single sensor the stability region for stabilizing an unstable LTI plant can be increased in some situations, and also in applications towards control of unmanned aerial vehicles (UAVS) [18].

In this note, we study remote estimation using relay nodes, and investigate what information the individual relays should send to the gateway/fusion center. In a related setup considered in [13], the relay could only perform network coding that linearly combines two of the sensor transmissions using an XOR operation [10]. Here, we allow for the possibility of the relay combining multiple sensor transmissions using XOR operations [19], as well as the possibility of the relay forwarding the sensors' transmissions, which can give better performance in certain situations. The current work extends our recent contribution [1], which considered the situation with two sensors and one relay, and i.i.d. packet dropping links. In this note we generalize [1] to multiple sensors and relays, and additionally

consider time-varying packet reception probabilities governed by time-varying fading channel gains and the transmission power over this channel.

II. SYSTEM MODEL

Throughout this note, k will denote the discrete time index, i will be used to denote the sensor indices, and l the relay index. The process is a discrete time linear system

$$x_{k+1} = Ax_k + w_k \quad (1)$$

where $x_k \in \mathbb{R}^n$ and $\{w_k\}$ is i.i.d. Gaussian with zero mean and covariance matrix $Q > 0$.¹ The process is observed by M sensors with measurements

$$y_{i,k} = C_i x_k + v_{i,k}, \quad i = 1, \dots, M \quad (2)$$

with $y_{i,k} \in \mathbb{R}$, $\forall i$ and $\{v_{i,k}\}$ are i.i.d. Gaussian with zero mean and variance $R_i \geq 0$, $i = 1, \dots, M$. The processes $\{v_{i,k}\}$ and $\{w_k\}$ are assumed to be mutually independent, with (A, C) detectable and $(A, Q^{1/2})$ stabilizable, where $C \triangleq \text{col}(C_1, \dots, C_M)$.

We assume that the measurements $y_{i,k}$ have undergone source coding and can be grouped into packets of b bits, with each packet short enough to be transmitted within one time step. In particular, the uniform quantizer of [20] will be used here. Under the additive noise model for quantization (which in general is quite accurate for bit rates as low as three bits per sample [21]), the quantized value of $y_{i,k}$ can be written as

$$y_{i,k}^q = y_{i,k} + q_{i,k}$$

where the quantization noise $q_{i,k}$ has variance $\delta_b \mathbb{E}[y_{i,k}^2]$, with

$$\delta_b = \frac{4b \ln 2}{3 \times 2^{2b}}$$

when using the uniform quantizer of [20]. The measurements $y_{i,k}^q$ are transmitted over orthogonal/parallel channels to a gateway, which will perform the Kalman filtering operation. Let $\gamma_{i,k}$, for $i = 1, \dots, M$, be random variables such that $\gamma_{i,k} = 1$ if $y_{i,k}^q$ is successfully transmitted to the gateway by sensor i , and $\gamma_{i,k} = 0$ otherwise.

Furthermore, there exist L intermediate relay nodes that can be used to aid the transmission of the sensor measurements to the gateway. Such situations can for instance occur in mesh networks, where nodes close to the process will make measurements of the process, while the other nodes don't make measurements but can be used to relay the sensor measurements to the gateway [14]. A diagram of the system model for the case of $M = 2$ sensors and $L = 2$ relays is given in Fig. 1. Each relay can listen to a subset of the sensor transmissions. Denote $\mathcal{I} = \{1, \dots, M\}$ as the set of all sensors, and $\mathcal{I}_l \subseteq \mathcal{I}$ as the set of sensors which relay l can listen to. In general the sets \mathcal{I}_l , $l = 1, \dots, L$ will not necessarily be disjoint, with possibly multiple relays listening to a given sensor. For $i \in \mathcal{I}_l$, let $\zeta_{i,k}^l$ be a random variable such that $\zeta_{i,k}^l = 1$ if the transmission at time k of sensor i is received by relay l , and $\zeta_{i,k}^l = 0$ otherwise. The relays can perform some simple local processing before transmitting over orthogonal channels to the gateway. Let $\tilde{\gamma}_{l,k}$ for $l = 1, \dots, L$ be random variables such that $\tilde{\gamma}_{l,k} = 1$ if transmission at time k from relay l to the gateway is successful, and $\tilde{\gamma}_{l,k} = 0$ otherwise.

¹We say that a matrix $X > 0$, if it is positive definite, and $X \geq 0$, if it is positive semi-definite.

This work was supported by the Australian Research Council. A preliminary version of parts of this note was presented in [1].

A. Leong is with the Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Vic. 3010, Australia, asleong@unimelb.edu.au. D. Quevedo is with the Faculty of Electrical Engineering and Information Technology (EIM-E), University of Paderborn, Paderborn, Germany, dquevedo@ieee.org.

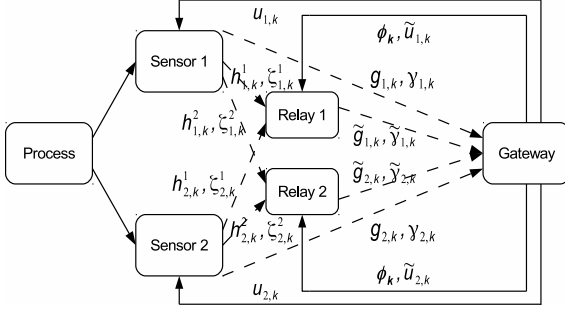


Fig. 1. System model for the case of two sensors and two relays

In this note we will consider a few simple operations that the relay can perform.² A relay can either: a) listen to one of the sensors' transmissions, say sensor i , and forward $y_{i,k}^q$ if it is successfully received by the relay, or b) listen to a number of the sensors' transmissions, say sensors i_1, i_2, \dots, i_L , and send $y_{i_1,k}^q \oplus y_{i_2,k}^q \oplus \dots \oplus y_{i_L,k}^q$ if $y_{i_1,k}^q, y_{i_2,k}^q, \dots, y_{i_L,k}^q$ have all been successfully received by the relay, where \oplus is the XOR operation. The XOR operation is commonly used in network coding [10], [19]. For instance, if the gateway receives both $y_{i,k}^q$ and $y_{i,k}^q \oplus y_{j,k}^q$, or both $y_{j,k}^q$ and $y_{i,k}^q \oplus y_{j,k}^q$, then the gateway can recover both $y_{i,k}^q$ and $y_{j,k}^q$ with a binary subtraction. In general, given the transmissions received at the gateway, the measurements which can be recovered can be determined using Gaussian elimination over \mathbb{Z}_2 .³ Determining which sensor/s each relay listens to, and which operation each relay uses, is one of the key questions addressed in this work, see Section V. We define a *relay configuration*

$$\phi_k = (\phi_{1,k}, \dots, \phi_{L,k})$$

at time k as the set of operations $\phi_{l,k}$ that each relay uses at time k . The set of all possible relay configurations will be denoted by Φ .

The communication channels will be modelled as time-varying fading channels [23]. We let $g_{i,k}, i = 1, \dots, M$ be the channel gains at time k from the sensor i to the gateway, $\tilde{g}_{l,k}, l = 1, \dots, L$ the channels gains from relay l to the gateway, and $h_{i,l,k}, i \in \mathcal{I}_l, l = 1, \dots, M$ the channel gains from sensor i to relay l . We use the block fading model [24] and assume that $\{g_{i,k}\}, \{\tilde{g}_{l,k}\}, \{h_{i,l,k}\}$ vary over time k in an i.i.d. manner, with the processes being mutually independent. Denote the transmit powers at time k of the sensors and relays by $u_{i,k}, i = 1, \dots, M$ and $\tilde{u}_{l,k}, l = 1, \dots, L$ respectively. Following the model of [9], the packet reception probabilities will depend on both the channel gains and transmit powers as follows: We have $\lambda_{i,k} \triangleq \mathbb{P}(\gamma_{i,k} = 1 | g_{i,k}, u_{i,k}) = p(g_{i,k} u_{i,k})$ as the time-varying⁴ packet reception probabilities from sensor i to the gateway, $\tilde{\lambda}_{l,k} \triangleq \mathbb{P}(\tilde{\gamma}_{l,k} = 1 | \tilde{g}_{l,k}, \tilde{u}_{l,k}) = p(\tilde{g}_{l,k} \tilde{u}_{l,k})$ the probabilities from relay l to the gateway, and $\rho_{i,l,k}^l \triangleq \mathbb{P}(\zeta_{i,l,k}^l = 1 | h_{i,l,k}, u_{i,k}) = p(h_{i,l,k}^l u_{i,k})$ the probabilities from sensor i to relay l . Here $p(\cdot) : [0, \infty) \rightarrow [0, 1]$ is a continuous monotonically increasing function whose form depends on the particular digital modulation and coding scheme being used [23]. For example, in the case of uncoded binary phase shift keying

²We assume limited computational power at the relays, thus only simple operations at a bit-level are considered in this note. If, however, additional computational capability is available, then other possibilities include the use of more involved network coding schemes [11] or the computation of local state estimates at the relays [22].

³While our use of the XOR operation is similar to network coding, our objectives are not exactly the same. In network coding transmissions are often regarded as "successful" only if all packets arrive (eventually) at their intended destinations, whereas in our problem even if some packets are not received one still can perform state estimation using the available measurements.

⁴This differs from the i.i.d. Bernoulli packet drop model with constant packet reception probabilities considered in e.g. [2]

TABLE I
NOTATION FOR DIFFERENT TYPES OF LINKS

	Channel Gain	Packet Reception Random Variable	Packet Reception Probability
Sensor i to Gateway	$g_{i,k}$	$\gamma_{i,k}$	$\lambda_{i,k}$
Relay l to Gateway	$\tilde{g}_{l,k}$	$\tilde{\gamma}_{l,k}$	$\tilde{\lambda}_{l,k}$
Sensor i to Relay l	$h_{i,l,k}^l$	$\zeta_{i,l,k}^l$	$\rho_{i,l,k}^l$

TABLE II
BOOLEAN EXPRESSIONS FOR $\theta_{1,k}$ AND $\theta_{2,k}$, FOR TWO SENSORS AND ONE RELAY, UNDER THREE TYPES OF OPERATIONS

ϕ_k	Forward $y_{1,k}^q$	Forward $y_{2,k}^q$	Send $y_{1,k}^q \oplus y_{2,k}^q$
$\theta_{1,k}$	$\gamma_{1,k} \vee (\tilde{\gamma}_{1,k} \wedge \zeta_{1,k}^1)$	$\gamma_{1,k}$	$\gamma_{1,k} \vee (\tilde{\gamma}_{1,k} \wedge \gamma_{2,k} \wedge \zeta_{1,k}^1 \wedge \zeta_{2,k}^1)$
$\theta_{2,k}$	$\gamma_{2,k}$	$\gamma_{2,k} \vee (\tilde{\gamma}_{1,k} \wedge \zeta_{2,k}^1)$	$\gamma_{2,k} \vee (\tilde{\gamma}_{1,k} \wedge \gamma_{1,k} \wedge \zeta_{1,k}^1 \wedge \zeta_{2,k}^1)$

(BPSK) transmission with b bits per packet, $p(\cdot)$ would take the form

$$p(gu) = \left(\int_{-\infty}^{\sqrt{gu}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \right)^b \quad (3)$$

where we assume a packet is successfully received if all b bits are successfully received. However, if there is channel coding and/or different digital modulation schemes, $p(\cdot)$ will take on different forms [25]. In Table I we summarize the notation for the channel gains, packet reception random variables and packet reception probabilities for the different types of links.

In addition, there are feedback links from the gateway to the sensors and relays which can be used to communicate the relay configuration ϕ_k and power levels $u_{i,k}$ and $\tilde{u}_{l,k}$ to be used, see Section V and Section VII. In this note we will assume that transmissions can occur over a much faster time scale than the process (1). Thus, delays experienced by the measurements in passing through intermediate relay nodes will be ignored. For instance, in the industrial wireless sensor networks standard WirelessHART [26], transmissions between nodes would typically take around 10 ms, whereas for many estimation and control applications the process time constant might be 1 sec or more.

III. KALMAN FILTER WITH PACKET DROPS AND RELAYS

Let $\theta_{i,k}, i = 1, \dots, M$ be random variables such that $\theta_{i,k} = 1$ if $y_{i,k}^q$ can be reconstructed at the gateway, and $\theta_{i,k} = 0$ otherwise. Note that in general $\theta_{i,k}$ and $\theta_{j,k}, i \neq j$ are not independent. Values of $\theta_{i,k}$ for different relay configurations and combinations of $\gamma_{i,k}, \tilde{\gamma}_{l,k}, \zeta_{i,l,k}^l$, can be written in Boolean algebra form. For example, in Table II we give the Boolean expressions for $\theta_{1,k}$ and $\theta_{2,k}$ in the case of two sensors and one relay, where we use the notation \wedge to denote logical 'and' and \vee to denote logical 'or'. Now define $\check{C}_k \triangleq \text{col}(\theta_{1,k} C_1, \dots, \theta_{M,k} C_M)$, $\check{y}_k \triangleq \text{col}(\theta_{1,k} y_{1,k}^q, \dots, \theta_{M,k} y_{M,k}^q)$, $\hat{x}_{k+1|k} \triangleq \mathbb{E}[x_{k+1} | \check{y}_0, \dots, \check{y}_k, \check{C}_0, \dots, \check{C}_k]$, $P_{k+1|k} \triangleq \mathbb{E}[(x_{k+1} - \hat{x}_{k+1|k})(x_{k+1} - \hat{x}_{k+1|k})^T | \check{y}_0, \dots, \check{y}_k, \check{C}_0, \dots, \check{C}_k]$. (4)

The associated Kalman filter equations which are run at the gateway can be written as (see e.g. [12])

$$\begin{aligned} \hat{x}_{k+1|k} &= A \hat{x}_{k|k-1} + K_k (\check{y}_k - \check{C}_k \hat{x}_{k|k-1}) \\ P_{k+1|k} &= A P_{k|k-1} A^T + Q - K_k \check{C}_k P_{k|k-1} A^T \end{aligned} \quad (5)$$

where $K_k = A P_{k|k-1} \check{C}_k^T (\check{C}_k P_{k|k-1} \check{C}_k^T + \check{R}_k)^{-1}$, with $\check{R}_k = \text{diag}(\check{R}_{1,k}, \dots, \check{R}_{M,k}) \triangleq \text{diag}(R_1 + \delta_b \mathbb{E}[y_{1,k}^2], \dots, R_M + \delta_b \mathbb{E}[y_{M,k}^2])$, similar to [13]. In the sequel, we will also call $P_k \triangleq P_{k|k-1}$.

Remark III.1. As in [13], the Kalman filter (5) uses all successfully reconstructed measurements, but those measurements where $\theta_{i,k} = 0$ are not taken into account.

IV. PERFORMANCE OF THE KALMAN FILTER WITH RELAYS

In Section II we have proposed that each relay can either listen to transmissions from one of the sensors which it then forwards to the gateway, or it listens to a number of sensors and performs an XOR operation that is then sent to the gateway. We wish to investigate which operation each relay should use, and which sensors each relay should listen to, i.e. determining the relay configuration ϕ_k , in order to give the best performance for the Kalman filter. This section presents some preliminary results on the performance of the Kalman filter, with optimal relay configuration selection to be studied in Section V. For time-varying Kalman filters with packet drops, one commonly used performance measure is the expected error covariance. In this note we consider the problem of optimal relay configuration selection in order to minimize the one step ahead expected error covariance $\mathbb{E}[P_{k+1}|P_k, \mathbf{g}_k, \phi_k]$, where

$$\mathbf{g}_k \triangleq \{g_{1,k}, \dots, g_{M,k}, \tilde{g}_{1,k}, \dots, \tilde{g}_{L,k}, h_{1,k}^1, \dots, h_{M,k}^L\} \quad (6)$$

represents the channel gains at time k , which in turn will determine the packet reception probabilities $\lambda_{i,k}, \tilde{\lambda}_{l,k}, \rho_{i,k}^l, i = 1, \dots, M, l = 1, \dots, L$. In order to compute $\mathbb{E}[P_{k+1}|P_k, \mathbf{g}_k, \phi_k]$, we will further assume that full channel state information (CSI) at the receiver is available, so that \mathbf{g}_k is known at the gateway.⁵ Define

$$f_k(X) \triangleq X A X^T + Q - \mathbb{E} \left[A X \check{C}_k^T \left(\check{C}_k X \check{C}_k^T + \check{R}_k \right)^{-1} \check{C}_k X A^T \middle| \mathbf{g}_k, \phi_k \right] \quad (7)$$

where the expectation is with respect to $\theta_{1,k}, \dots, \theta_{M,k}$ in the definition of \check{C}_k in (4). Equivalently, we can write $f_k(X)$ as:

$$f_k(X) = A X A^T + Q - \sum_{I \subseteq \mathcal{I}} \mathbb{E} \left[\prod_{i \in I} \Theta_{i,k} \prod_{j \notin I} (1 - \Theta_{j,k}) \middle| \mathbf{g}_k, \phi_k \right] A X \bar{C}(I)^T \times \left(\bar{C}(I) X \bar{C}(I)^T + \bar{R}_k(I) \right)^{-1} \bar{C}(I) X A^T \quad (8)$$

where $\bar{C}(I) = \text{col}(\{C_i, i \in I\})$, $\bar{R}_k(I) = \text{diag}(\{\check{R}_{i,k}, i \in I\})$, and $\Theta_{i,k}, i = 1, \dots, M$ are random variables with the same distributions as $\theta_{i,k}$. The quantities $\mathbb{E} \left[\prod_{i \in I} \Theta_{i,k} \prod_{j \notin I} (1 - \Theta_{j,k}) \middle| \mathbf{g}_k, \phi_k \right]$ can be computed in terms of the packet reception probabilities $\lambda_{i,k}, \tilde{\lambda}_{l,k}, \rho_{i,k}^l, i = 1, \dots, M, l = 1, \dots, L$. A systematic procedure for doing this is as follows:

1) Write out the Boolean expression

$$\bigwedge_{i \in I} \theta_{i,k} \bigwedge_{j \notin I} (\neg \theta_{j,k}) \quad (9)$$

where each $\theta_{i,k}$ is written as a Boolean expression, $\neg \theta_{j,k}$ denotes the negation of the Boolean expression for $\theta_{j,k}$, and the notation $\bigwedge_{i \in I} \theta_{i,k} \triangleq \theta_{i_1,k} \wedge \dots \wedge \theta_{i_n,k} \wedge \dots$ for indices $i_n \in I$.

2) Convert the Boolean expression (9) into the sum of products normal form [27]. (Note that this can be done systematically.)

3) $\mathbb{E} \left[\prod_{i \in I} \Theta_{i,k} \prod_{j \notin I} (1 - \Theta_{j,k}) \middle| \mathbf{g}_k, \phi_k \right]$ is then given by taking the sum of products normal form of (9), and replacing \wedge with multiplication, \vee with addition, $\gamma_{i,k}$ with $\lambda_{i,k}$, $\neg \gamma_{i,k}$ with $1 - \lambda_{i,k}$, $\tilde{\gamma}_{i,k}$ with $\tilde{\lambda}_{i,k}$, $\neg \tilde{\gamma}_{i,k}$ with $1 - \tilde{\lambda}_{i,k}$, $\zeta_{i,k}^l$ with $\rho_{i,k}^l$, and $\neg \zeta_{i,k}^l$ with $1 - \rho_{i,k}^l$.

By step 2), each term in the sum will correspond to a distinct entry of the truth table for $\theta_{1,k}, \dots, \theta_{M,k}$, thus allowing $\mathbb{E} \left[\prod_{i \in I} \Theta_{i,k} \prod_{j \notin I} (1 - \Theta_{j,k}) \middle| \mathbf{g}_k, \phi_k \right]$ to be easily calculated.

We now give a result on how the packet reception probabilities affect the expected error covariance $\mathbb{E}[P_{k+1}|P_k, \mathbf{g}_k, \phi_k] = f_k(P_k)$. First denote:

$$\begin{aligned} \mathcal{L}_{i,k} &\triangleq \{\lambda_{1,k}, \dots, \lambda_{i,k}, \dots, \lambda_{M,k}, \tilde{\lambda}_{1,k}, \dots, \tilde{\lambda}_{L,k}, \rho_{1,k}^1, \dots, \rho_{M,k}^L\} \\ \mathcal{U}_{i,k} &\triangleq \{\lambda_{1,k}, \dots, \mu_{i,k}, \dots, \lambda_{M,k}, \tilde{\lambda}_{1,k}, \dots, \tilde{\lambda}_{L,k}, \rho_{1,k}^1, \dots, \rho_{M,k}^L\} \end{aligned}$$

⁵In practice this can be achieved using channel estimation and prediction algorithms, see references in [9], [13].

$$\begin{aligned} \tilde{\mathcal{L}}_{l,k} &\triangleq \{\lambda_{1,k}, \dots, \lambda_{M,k}, \tilde{\lambda}_{1,k}, \dots, \tilde{\lambda}_{l,k}, \dots, \tilde{\lambda}_{L,k}, \rho_{1,k}^1, \dots, \rho_{M,k}^L\} \\ \tilde{\mathcal{U}}_{l,k} &\triangleq \{\lambda_{1,k}, \dots, \lambda_{M,k}, \tilde{\lambda}_{1,k}, \dots, \tilde{\mu}_{l,k}, \dots, \tilde{\lambda}_{L,k}, \rho_{1,k}^1, \dots, \rho_{M,k}^L\} \\ \mathcal{R}_{i,k}^l &\triangleq \{\lambda_{1,k}, \dots, \lambda_{M,k}, \tilde{\lambda}_{1,k}, \dots, \tilde{\lambda}_{L,k}, \rho_{1,k}^1, \dots, \rho_{i,k}^l, \dots, \rho_{M,k}^L\} \\ \mathcal{S}_{i,k}^l &\triangleq \{\lambda_{1,k}, \dots, \lambda_{M,k}, \tilde{\lambda}_{1,k}, \dots, \tilde{\lambda}_{L,k}, \rho_{1,k}^1, \dots, \sigma_{i,k}^l, \dots, \rho_{M,k}^L\} \end{aligned}$$

Lemma IV.1. Let $f_{\mathcal{X}_{i,k}}(\cdot)$ be defined by $f_k(\cdot)$ in (7) when the links have packet reception probabilities $\mathcal{X}_{i,k}$. Then, irrespective of which relay configuration is used, $\forall i = 1, \dots, M, \forall l = 1, \dots, L$,

$$\begin{aligned} \lambda_{i,k} \leq \mu_{i,k} &\Rightarrow f_{\mathcal{L}_{i,k}}(X) \geq f_{\mathcal{U}_{i,k}}(X) \\ \tilde{\lambda}_{l,k} \leq \tilde{\mu}_{l,k} &\Rightarrow f_{\tilde{\mathcal{L}}_{l,k}}(X) \geq f_{\tilde{\mathcal{U}}_{l,k}}(X) \\ \rho_{i,k}^l \leq \sigma_{i,k}^l &\Rightarrow f_{\mathcal{R}_{i,k}^l}(X) \geq f_{\mathcal{S}_{i,k}^l}(X) \end{aligned}$$

Proof: Consider the case $\lambda_{i,k} \leq \mu_{i,k}$. Recall that Bernoulli random variables can be generated from $U(0,1)$ uniform random variables, by comparing the uniform random variable with the probability that the Bernoulli random variable is equal to one, i.e.. $\gamma_{i,k} = 1$ when $u \leq \lambda_{i,k}$, and $\gamma_{i,k} = 0$ otherwise, where u is $U(0,1)$. Let ω denote an outcome corresponding to N independent realizations of $U(0,1)$ random variables, where N is equal to the total number of packet dropping links. For each ω , one can generate corresponding independent Bernoulli random variables $\gamma_{1,k}, \dots, \gamma_{M,k}, \tilde{\gamma}_{1,k}, \dots, \tilde{\gamma}_{L,k}, \zeta_{1,k}^1, \dots, \zeta_{M,k}^L$. One can then construct the Bernoulli random variables $\theta_{1,k}, \dots, \theta_{M,k}$, and hence \check{C}_k as in (4).

Let $\check{C}_{\mathcal{L}_{i,k}}(\omega)$ be the matrix \check{C}_k when using packet reception probabilities $\mathcal{L}_{i,k}$ and $\check{C}_{\mathcal{U}_{i,k}}(\omega)$ be the matrix \check{C}_k when using packet reception probabilities $\mathcal{U}_{i,k}$. Now note that if $\theta_{j,k}(\omega) = 1$ using the packet reception probabilities $\mathcal{L}_{i,k}$, then we also have $\theta_{j,k}(\omega) = 1$ when using the packet reception probabilities $\mathcal{U}_{i,k}$, by the way in which $\theta_{j,k}(\omega)$ is constructed, and since an increase in the packet reception probability of any link cannot decrease the probability of reconstructing any of the sensor measurements. Hence

$$\begin{aligned} A X \check{C}_{\mathcal{L}_{i,k}}(\omega)^T \left(\check{C}_{\mathcal{L}_{i,k}}(\omega) X \check{C}_{\mathcal{L}_{i,k}}(\omega)^T + \check{R}_k \right)^{-1} \check{C}_{\mathcal{L}_{i,k}}(\omega) X A^T \\ \geq A X \check{C}_{\mathcal{U}_{i,k}}(\omega)^T \left(\check{C}_{\mathcal{U}_{i,k}}(\omega) X \check{C}_{\mathcal{U}_{i,k}}(\omega)^T + \check{R}_k \right)^{-1} \check{C}_{\mathcal{U}_{i,k}}(\omega) X A^T \end{aligned} \quad (10)$$

Since (10) holds for all ω , we have

$$\begin{aligned} \mathbb{E} \left[A X (\check{C}_{\mathcal{L}_{i,k}})^T \left(\check{C}_{\mathcal{L}_{i,k}} X (\check{C}_{\mathcal{L}_{i,k}})^T + \check{R}_k \right)^{-1} \check{C}_{\mathcal{L}_{i,k}} X A^T \middle| \mathbf{g}_k, \phi_k \right] \\ \geq \mathbb{E} \left[A X (\check{C}_{\mathcal{U}_{i,k}})^T \left(\check{C}_{\mathcal{U}_{i,k}} X (\check{C}_{\mathcal{U}_{i,k}})^T + \check{R}_k \right)^{-1} \check{C}_{\mathcal{U}_{i,k}} X A^T \middle| \mathbf{g}_k, \phi_k \right] \end{aligned}$$

which shows that $f_{\mathcal{L}_{i,k}}(X) \geq f_{\mathcal{U}_{i,k}}(X)$. The other two cases can be proved in a similar manner. ■

Lemma IV.1 says that increasing the packet arrival rate on any one of the packet dropping links will give a decrease in the value of the one step ahead expected error covariance, no matter which relay configuration is used.

We also wish to determine how the system parameters such as $C_1, \dots, C_M, R_1, \dots, R_M$ will affect the values of the expected error covariance. Denote:

$$\begin{aligned} \mathcal{A}_i &\triangleq \{C_1, \dots, C_i, \dots, C_M, R_1, \dots, R_i, \dots, R_M\} \\ \mathcal{B}_i &\triangleq \{C_1, \dots, D_i, \dots, C_M, R_1, \dots, S_i, \dots, R_M\} \end{aligned}$$

Lemma IV.2. Let $f_{\mathcal{Y}_{i,k}}(X)$ be defined by $f_k(\cdot)$ in (8) when the sensors have parameters \mathcal{Y}_i . Then, irrespective of which relay configuration is used, $\forall i = 1, \dots, M$,

$$C_i^T \check{R}_{i,k}^{-1} C_i \leq D_i^T \check{S}_{i,k}^{-1} D_i \Rightarrow f_{\mathcal{A}_i}(X) \geq f_{\mathcal{B}_i}(X)$$

The proof of Lemma IV.2 is omitted; a related result is proved in [1]. The quantity $C_i^T \check{R}_{i,k}^{-1} C_i$ can be regarded as the signal-to-noise ratio (SNR) for sensor i (which includes the quantization

noise contribution). Thus Lemma IV.2 shows that the expected error covariance is monotonic in the SNRs $C_i^T \check{R}_{i,k}^{-1} C_i$.

V. OPTIMAL RELAY CONFIGURATION SELECTION

We now wish to address the question of determining which configurations for the relays will give the best Kalman filter performance. Suppose for now that the sensor transmit powers $u_{i,k}$, $i = 1, \dots, M$ and relay transmit powers $\tilde{u}_{l,k}$, $l = 1, \dots, L$ are given or fixed (The more difficult problem of jointly optimizing the relay configuration and transmission powers will be considered in Section VII.). We wish to choose at each time instant k , the relay configuration ϕ_k^* that minimizes $\text{Tr}\mathbb{E}[P_{k+1}|P_k, \mathbf{g}_k, \phi_k]$, i.e.

$$\phi_k^* = \arg \min_{\phi_k \in \Phi} \text{Tr}\mathbb{E}[P_{k+1}|P_k, \mathbf{g}_k, \phi_k]. \quad (11)$$

where $\mathbb{E}[P_{k+1}|P_k, \mathbf{g}_k, \phi_k] = f_k(P_k)$, see (8).

A. Optimal relay configuration selection

Problem (11) can, in principle, be solved by exhaustive search at the gateway. The optimal configuration can then be fed back to the relays. We will characterize the number of relay configurations that need to be checked at each time instant for exhaustive search.

Lemma V.1. *Let \mathcal{I}_l be the set of sensors that relay l can listen to, and let $M_l = |\mathcal{I}_l|$ denote the cardinality of \mathcal{I}_l . Suppose that there are no restrictions on how many relays listen to the same sensor. Then the number of possible relay configurations for ϕ_k is*

$$|\Phi| = \prod_{l=1}^L (2^{M_l} - 1) \quad (12)$$

Proof: First fix a relay l , which can listen to M_l of the sensors. This relay can either forward any one of the sensor transmissions, or perform the XOR operation on two or more of the sensor transmissions it listens to, resulting in $M_l + \binom{M_l}{2} + \binom{M_l}{3} + \dots + \binom{M_l}{M_l} = 2^{M_l} - 1$ possible operations. If there are no restrictions on multiple relays listening to the same sensor, then by the multiplication principle the number of relay configurations is $\prod_{l=1}^L (2^{M_l} - 1)$. ■

We thus see that the number of configurations that needs to be checked is in the worst case (where each relay can listen to all sensors) exponential in the number of relays L and number of sensors M . However, in practice, due to geographical considerations, the number of sensors M_l that each sensor l listens to is often small, e.g. in [28], [29] it is assumed that $M_l \leq 3$.

B. Stability of Kalman filtering with relay configuration selection

We now wish to give a condition for stability of the Kalman filter with optimal relay configuration selection.

Definition 1. *The Kalman filter is said to be exponentially bounded if there exist finite constants α and β , and $\rho \in [0, 1)$, such that $\text{Tr}\mathbb{E}[P_k] \leq \alpha \rho^k + \beta$, $\forall k$.*

Theorem V.2. *Let $\{s_k\}$ be a stochastic process such that $s_k = 1$ if \check{C}_k is full rank, and $s_k = 0$ otherwise. Suppose there exists a policy $\phi^\sharp(\mathbf{g}_k)$ dependent only on \mathbf{g}_k , such that*

$$\|A\|^2 \mathbb{E}[\mathbb{P}(s_k = 0 | \mathbf{g}_k, \phi^\sharp(\mathbf{g}_k))] < 1, \quad (13)$$

where $\|A\|$ is the spectral norm of A . Then the Kalman filter using the configurations obtained from problem (11) is exponentially bounded.

Proof: Since the distribution of \check{C}_k depends on P_k and \mathbf{g}_k , and \mathbf{g}_k is independent in time and of P_k , we have $\mathbb{P}(\check{C}_k | P_k, P_{k-1}, \dots, P_0) = \mathbb{P}(\check{C}_k | P_k)$. Then by (5), the process $\{P_k\}$ is Markovian. Now define $V_k \triangleq \text{Tr}P_k$. Since in

the relay configuration selection problem (11) we are minimizing $\mathbb{E}\{V_{k+1} | P_k, \mathbf{g}_k, \phi_k(P_k, \mathbf{g}_k)\}$, we have

$$\begin{aligned} \mathbb{E}\{V_{k+1} | P_k\} &= \mathbb{E}[\mathbb{E}\{V_{k+1} | P_k, \mathbf{g}_k, \phi_k^*(P_k, \mathbf{g}_k)\}] \\ &\leq \mathbb{E}[\mathbb{E}\{V_{k+1} | P_k, \mathbf{g}_k, \phi^\sharp(\mathbf{g}_k)\}] \\ &= \mathbb{E}[\mathbb{E}\{V_{k+1} | P_k, \mathbf{g}_k, \phi^\sharp(\mathbf{g}_k), s_k = 1\} \mathbb{P}\{s_k = 1 | P_k, \mathbf{g}_k, \phi^\sharp(\mathbf{g}_k)\} \\ &\quad + \mathbb{E}\{V_{k+1} | P_k, \mathbf{g}_k, \phi^\sharp(\mathbf{g}_k), s_k = 0\} \mathbb{P}\{s_k = 0 | P_k, \mathbf{g}_k, \phi^\sharp(\mathbf{g}_k)\}] \\ &\leq W + (\|A\|^2 V_k + \text{Tr}Q) \mathbb{E}[\mathbb{P}\{s_k = 0 | \mathbf{g}_k, \phi^\sharp(\mathbf{g}_k)\}] \end{aligned}$$

where the last inequality is shown using similar arguments to [9], [13], and W is a positive constant. If $\|A\|^2 \mathbb{E}[\mathbb{P}(s_k = 0 | \mathbf{g}_k, \phi^\sharp(\mathbf{g}_k))] < 1$ we may then use a stochastic Lyapunov function argument similar to [13] to show that $\mathbb{E}\{V_k | P_0\} \leq \alpha \rho^k + \beta$, $\forall k$ for some $\rho \in [0, 1)$ and constants α and β , which establishes exponential boundedness of the Kalman filter. ■

Theorem V.2 thus provides a sufficient condition for Kalman filter stability dependent on the system matrix A and the distributions of the channel gains \mathbf{g}_k .

Example V.3. *Consider the case of two sensors and one relay, with \check{C}_k being full rank only when $\theta_{1,k} = \theta_{2,k} = 1$. Then*

$$\begin{aligned} \mathbb{P}(\theta_{1,k} = 1, \theta_{2,k} = 1 | \mathbf{g}_k, \phi_k) &= \mathbb{E}[\Theta_{1,k} \Theta_{2,k} | \mathbf{g}_k, \phi_k] \\ &= \lambda_{1,k} \lambda_{2,k} + (1 - \lambda_{1,k}) \lambda_{2,k} \tilde{\lambda}_{1,k} \rho_{1,k}^1 \end{aligned}$$

Suppose we choose ϕ^\sharp to be the suboptimal policy that always forwards $y_{1,k}^q$, and with the transmit powers $u_{1,k} = u_{2,k} = \tilde{u}_{1,k} = 1$. The condition (13) then becomes

$$\begin{aligned} \mathbb{E}[\mathbb{P}(s_k = 0 | \mathbf{g}_k, \phi^\sharp(\mathbf{g}_k))] &= \int \left((1 - p(g_{1,k})p(g_{2,k}) \right. \\ &\quad \left. - (1 - p(g_{1,k}))p(g_{2,k})p(\tilde{g}_{1,k})p(h_{1,k}^1) \right) d\mathbb{P}(\mathbf{g}_k) < \frac{1}{\|A\|^2} \end{aligned}$$

which can be checked by numerically computing the integral for specific functions $p(\cdot)$ and probability distributions $\mathbb{P}(\mathbf{g}_k)$. If condition (13) is satisfied for this suboptimal policy, then by Theorem V.2 the Kalman filter using the optimal relay configurations will also be stable.

C. Suboptimal relay configuration selection

Lemma V.1 has shown that the optimal way of choosing the relay configuration by checking each configuration is exponential in the number of relays L , which is computationally intensive when L is large. To reduce computational complexity, a suboptimal method for determining a relay configuration is to optimize the operation of each relay l independently of each other. The motivation for this method is that sometimes other relays may become unavailable, thus one should optimize the performance of each relay irrespective of what the other relays are doing. Specifically, consider subsets $I_l \subseteq \mathcal{I}_l$. Let $\bar{C}(I_l) = \text{col}(\{C_i, i \in I_l\})$, $\bar{R}_k(I_l) = \text{diag}(\{\check{R}_{i,k}, i \in I_l\})$, and

$$\begin{aligned} f_{l,k}(X) &\triangleq AXA^T + Q - \sum_{I_l \subseteq \mathcal{I}_l} \mathbb{E}^l \left[\prod_{i \in I_l} \Theta_{i,k} \prod_{j \notin I_l} (1 - \Theta_{j,k}) | \mathbf{g}_k, \phi_{l,k} \right] \\ &\quad \times AX\bar{C}(I_l)^T \left(\bar{C}(I_l)X\bar{C}(I_l)^T + \bar{R}_k(I_l) \right)^{-1} \bar{C}(I_l)XA^T \end{aligned}$$

where the terms $\mathbb{E}^l \left[\prod_{i \in I_l} \Theta_{i,k} \prod_{j \notin I_l} (1 - \Theta_{j,k}) | \mathbf{g}_k, \phi_{l,k} \right]$ are computed assuming that relay l is the only relay available. One then computes $f_{l,k}(P_k)$ for each of the operations $\phi_{l,k}$ that relay l can perform, with the one that gives the smallest value of $\text{Tr}(f_{l,k}(P_k))$ then chosen. This optimization can be carried out for each relay independently of the other relays. The number of configurations that need to be checked at each time step k is then $\sum_{l=1}^L (2^{M_l} - 1)$, which (compare to (12)) is no longer exponential in the number of relays L , and with M_l usually being small as mentioned at the end of Section V-A.

VI. A SPECIAL CASE

Here we consider a special case where additional analytical results can be derived. We will study the effects of varying the packet reception probabilities and signal-to-noise ratios, which will provide some insight into the general behaviour.

Recall that in the case of two sensors and one relay, the relay can either i) forward $y_{1,k}$ if it is received, ii) forward $y_{2,k}$ if it is received, or iii) send $y_{1,k}^q \oplus y_{2,k}^q$ if both $y_{1,k}^q$ and $y_{2,k}^q$ are received. We will consider a scalar example with $\lambda_{1,k} = \rho_{1,k}^1 = \rho_{2,k}^1 = 1, \forall k$, corresponding to the case where the channels from the relay to the gateway, and from the sensors to the relay, are error free. We will call $A = a, Q = q, C_1 = c_1, C_2 = c_2, R_1 = r_1, R_2 = r_2, \tilde{R}_{1,k} = \tilde{r}_{1,k}, \tilde{R}_{2,k} = \tilde{r}_{2,k}$. We have

$$\mathbb{E}[P_{k+1}|P_k, \mathbf{g}_k, \phi_k] = \begin{cases} a^2 P_k + q - \frac{\lambda_{2,k} a^2 P_k^2}{P_k + \frac{1}{c_1^2/\tilde{r}_{1,k} + c_2^2/\tilde{r}_{2,k}}} - \frac{(1-\lambda_{2,k})a^2 P_k^2}{P_k + \frac{1}{c_1^2/\tilde{r}_{1,k}}}, & \text{forward } y_{1,k}^q \\ a^2 P_k + q - \frac{\lambda_{1,k} a^2 P_k^2}{P_k + \frac{1}{c_1^2/\tilde{r}_{1,k} + c_2^2/\tilde{r}_{2,k}}} - \frac{(1-\lambda_{1,k})a^2 P_k^2}{P_k + \frac{1}{c_2^2/\tilde{r}_{2,k}}}, & \text{forward } y_{2,k}^q \\ a^2 P_k + q - \frac{(\lambda_{1,k} + \lambda_{2,k} - \lambda_{1,k}\lambda_{2,k})a^2 P_k^2}{P_k + \frac{1}{c_1^2/\tilde{r}_{1,k} + c_2^2/\tilde{r}_{2,k}}}, & \text{send } y_{1,k}^q \oplus y_{2,k}^q \end{cases} \quad (14)$$

We want to see under what conditions the XOR operation $y_{1,k}^q \oplus y_{2,k}^q$ outperforms forwarding of measurements. First let the terms $c_1^2/\tilde{r}_{1,k}$ and $c_2^2/\tilde{r}_{2,k}$, which can be regarded as the signal-to-noise ratios (SNRs) of sensors 1 and 2, be fixed. From (14) we see that sending $y_{1,k}^q \oplus y_{2,k}^q$ is better than forwarding $y_{1,k}^q$ when

$$\frac{\lambda_{1,k}(1-\lambda_{2,k})}{P_k + \frac{1}{c_1^2/\tilde{r}_{1,k} + c_2^2/\tilde{r}_{2,k}}} > \frac{1-\lambda_{2,k}}{P_k + \frac{1}{c_1^2/\tilde{r}_{1,k}}} \text{ or } \lambda_{1,k} > \frac{P_k + \frac{1}{c_1^2/\tilde{r}_{1,k} + c_2^2/\tilde{r}_{2,k}}}{P_k + \frac{1}{c_1^2/\tilde{r}_{1,k}}} \quad (15)$$

Similarly, sending $y_{1,k}^q \oplus y_{2,k}^q$ is better than forwarding $y_{2,k}^q$ when

$$\lambda_{2,k} > \frac{P_k + \frac{1}{c_1^2/\tilde{r}_{1,k} + c_2^2/\tilde{r}_{2,k}}}{P_k + \frac{1}{c_2^2/\tilde{r}_{2,k}}} \quad (16)$$

From (15)-(16), sending $y_{1,k}^q \oplus y_{2,k}^q$ is best when both packet reception probabilities $\lambda_{1,k}$ and $\lambda_{2,k}$ are above certain thresholds, which in turn implies that the instantaneous channel gains $g_{1,k}$ and $g_{2,k}$ need to be above some thresholds. Thus, for lower quality channels, forwarding of measurements gives better performance than the network coding operation. The intuitive explanation for this is that when the gateway receives $y_{1,k}^q \oplus y_{2,k}^q$, it needs one other measurement (either $y_{1,k}^q$ or $y_{2,k}^q$) in order to be useful. In contrast, if the relay forwards $y_{1,k}^q$ or $y_{2,k}^q$, this value (if received at the gateway) can be used even if the direct transmissions from the sensors are lost.

Alternatively, regard $\lambda_{1,k}$ and $\lambda_{2,k}$ as being fixed. Rewriting (15) and (16), sending $y_{1,k}^q \oplus y_{2,k}^q$ is better than forwarding $y_{1,k}^q$ when

$$\frac{\lambda_{1,k}}{c_1^2/\tilde{r}_{1,k}} - \frac{1}{c_1^2/\tilde{r}_{1,k} + c_2^2/\tilde{r}_{2,k}} > P_k(1-\lambda_{1,k}), \quad (17)$$

and sending $y_{1,k}^q \oplus y_{2,k}^q$ is better than forwarding $y_{2,k}^q$ when

$$\frac{\lambda_{2,k}}{c_2^2/\tilde{r}_{2,k}} - \frac{1}{c_1^2/\tilde{r}_{1,k} + c_2^2/\tilde{r}_{2,k}} > P_k(1-\lambda_{2,k}) \quad (18)$$

For fixed $\lambda_{1,k}$ and $\lambda_{2,k}$, we see that if either $c_1^2/\tilde{r}_{1,k}$ or $c_2^2/\tilde{r}_{2,k}$ is sufficiently large, then conditions (17) and (18) cannot both be simultaneously satisfied. The intuitive reason for requiring the signal-to-noise ratios to be small in order for network coding to give benefits, is that the relative performance gains by having both measurements available at the gateway (vs. just one of the measurements) is greater at low SNRs than high SNRs.

These qualitative observations, that the XOR operation needs channel conditions to be good, or for the signal-to-noise ratios to be low, in order to give benefits over forwarding of measurements,

have been observed in simulations for more general cases of packet dropping links to and from the relays, and with larger networks. Similar behaviour has also been reported in the network coding literature [29].

VII. RELAY CONFIGURATION SELECTION AND POWER CONTROL

In Section V the sensor and relay transmit powers were assumed to be fixed. However, the presence of time-varying fading channels will also allow for the use of power control techniques to further improve performance. In this section we present one possible formulation which optimizes the estimation performance subject to a sum of transmit powers constraint.

As in Section V, we assume that full channel state information (CSI) is available at the receiver, with \mathbf{g}_k in (6) representing the set of all channel gains at time k . The transmit powers of the sensors and relays can then depend on both the instantaneous channel gains \mathbf{g}_k and the error covariance P_k , with these transmit powers being computed at the gateway (which is assumed to have more computational resources than the sensors and relays) and fed back to the sensors and relays before transmission occurs at the next time step. Denote $\mathbf{u}_k(\mathbf{g}_k, P_k) \triangleq \{u_{1,k}(\mathbf{g}_k, P_k), \dots, u_{M,k}(\mathbf{g}_k, P_k), \tilde{u}_{1,k}(\mathbf{g}_k, P_k), \dots, \tilde{u}_{L,k}(\mathbf{g}_k, P_k)\}$ as the set of all transmit powers at time k .

A. Optimal power control for a given relay configuration

For a given relay configuration, we pose the following power control problem:

$$\begin{aligned} \min_{\mathbf{u}_k(\mathbf{g}_k, P_k)} \quad & \text{Tr}[\mathbb{E}[P_{k+1}|P_k, \mathbf{g}_k, \phi_k]] \\ \text{s.t.} \quad & \sum_{i=1}^M u_{i,k}(\mathbf{g}_k, P_k) + \sum_{l=1}^L \tilde{u}_{l,k}(\mathbf{g}_k, P_k) \leq u_{\text{tot}} \end{aligned} \quad (19)$$

which minimizes the expected one-step ahead error covariance subject to the sum power $\sum_{i=1}^M u_{i,k}(\mathbf{g}_k, P_k) + \sum_{l=1}^L \tilde{u}_{l,k}(\mathbf{g}_k, P_k)$ being less than u_{tot} . Due to the objective being a complicated nonlinear function of the transmit powers \mathbf{u}_k , the optimization problem (19) is in general non-convex and will need to be solved numerically using global optimization algorithms [30].

B. Joint relay configuration selection and power control

Problem (19) is for a given relay configuration. To optimally choose both the relay configuration and transmission powers, we can in principle solve $\prod_{l=1}^L (2^{M_l} - 1)$ instances of problem (19) at each time step (for each of the configurations, see Lemma V.1), and choose the relay configuration that gives the smallest value for the objective function, which however is very computationally intensive.

A less computationally intensive suboptimal scheme is to first choose a relay configuration by assuming a simple power allocation (e.g. that the total power u_{tot} is equally divided between the sensors and relays), and using the suboptimal method of Section V-C to choose a relay configuration. For this chosen relay configuration, we then further optimize the transmission powers by solving the power control problem (19).

VIII. NUMERICAL STUDIES

We first look at the performance differences between the optimal relay configuration selection and the suboptimal methods of Section V. We consider a situation with two sensors and two relays, where each of the relays can listen to both sensor transmissions, see Fig. 1. We consider the scalar case with $a = 0.95, q = 1, c_1 = c_2 = 1, r_1 = r_2 = 1$. For simplicity, we assume that the links from the sensors to the relays are perfect, with the fading channels (from the sensors to gateway, and from the relays to gateway) being exponentially distributed with mean 1, which models the case of

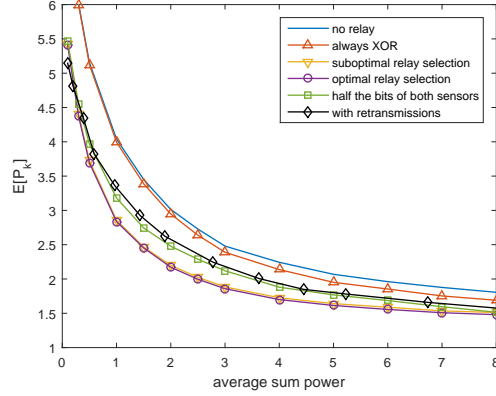


Fig. 2. Optimal and suboptimal relay configuration selection

Rayleigh fading [23]. Similar to [9], we assume that the digital communication uses BPSK transmission with $b = 6$ bits per packet, so that the function $p(\cdot)$ in Section II has the form (3). We distribute the transmit powers equally between the sensors and relays. Fig. 2 plots the average sum power and expected error covariance $\mathbb{E}[P_k]$ (obtained by time averaging $(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T$ over 10000 Monte Carlo iterations), for the optimal and suboptimal relay configuration selection methods. For comparison we also plot the performance for the cases of: 1) no relay, 2) a scheme where the relay always performs the XOR operation [13], 3) a scheme where the relay sends the two sensor measurements with less accuracy by removing half the bits⁶, and 4) a scheme where the gateway can ask for each lost transmission to be retransmitted once⁷. In each case, the expected error covariance decreases as the average power is increased. Since by (3) larger powers imply higher packet reception probabilities, this behaviour is in agreement with Lemma IV.1. We also see that the suboptimal method that optimizes each relay separately gives very close performance to the optimal method, and significantly outperforms the other schemes.

We next consider the case of two sensors and one relay, with Rayleigh fading for each of the fading channels. We choose $g_{1,k}, g_{2,k}$ to have mean 1, while $\tilde{g}_{1,k}, h_{1,k}, h_{2,k}$ have mean 4. This models the case where power decays in free space as $1/d^2$ where d is the distance from the transmitter [23], with the relay located approximately halfway between the sensors and gateway. In Fig. 3 we plot $\mathbb{E}[P_k]$ (obtained by time averaging $(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T$ over 10000 Monte Carlo iterations) for different sum powers, obtained by solving problem (19) using the `fmincon` routine in Matlab for each relay configuration and selecting the best one. We also plot the performance of the suboptimal scheme where a relay configuration is first chosen (assuming equal power allocation) and then power control is performed. We compare this with the case where there is no power control, with the sensors and relay using the same transmit power at all times, but with the best relay configuration chosen at each time step. Additionally, we plot the case where the relay always performs the XOR operation, and the case without a relay but with power control. We see that doing power control gives significant benefits, with the best performance when one optimizes both the relay configuration and transmit powers. The suboptimal scheme

⁶Here the relay removes from each quantized sensor measurement the least significant $b/2$ bits. Then in the Kalman filter equations (5), for the case where the direct transmissions are not successful, in the expression for \tilde{R}_k we replace $\delta_b = \frac{4b \ln 2}{3 \times 2^{2b}}$ with $\delta_{b/2} = \frac{2b \ln 2}{3 \times 2^b}$ whenever we use a measurement where half the bits have been removed.

⁷Here we assume that additional transmit power (same as the power for a single transmission) is used for each retransmission, with a successfully retransmitted measurement (from time k) available to the Kalman filter at time $k + 1$, which now utilizes a buffer similar to [7].

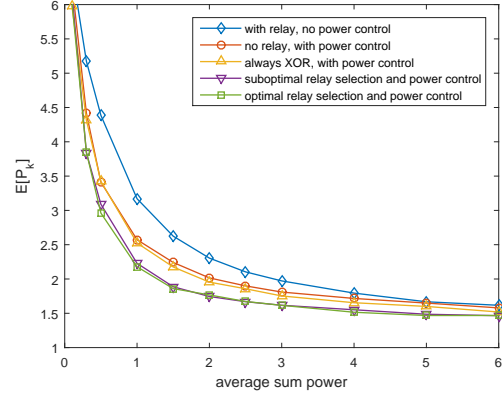


Fig. 3. Power control and relay configuration selection

where a relay configuration is first chosen by assuming equal power allocation, and then the powers are optimized, performs very close to the optimal scheme. Comparing the plots using power control, we see that for a given expected covariance the average power required is significantly less (at least 30-40%) when a relay is used.

IX. CONCLUSION

This note has studied the use of relays for Kalman filtering with multiple sensors over packet dropping links, where the packet reception probabilities are governed by fading channel gains and sensor and relay transmit powers. By allowing relays to either forward one of the sensor's measurements or perform a network coding operation, we have considered the problem of determining the optimal relay configuration at each time step, together with a simpler suboptimal method. We have also studied the use of power control in addition to selecting the best relay configuration, to further improve performance. Numerical results have demonstrated that the use of relays can lead to power savings of at least 30-40%. Future work will include studies of other strategies that relays can perform, and extending our setup to utilize relays in multi-hop networks [31].

REFERENCES

- [1] A. S. Leong and D. E. Quevedo, "On the use of a relay for Kalman filtering over packet dropping links," in *Proc. ACC*, Washington, DC, Jun. 2013.
- [2] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1453–1464, September 2004.
- [3] X. Liu and A. J. Goldsmith, "Kalman filtering with partial observation losses," in *Proc. IEEE Conf. Decision and Control*, Bahamas, Dec. 2004, pp. 1413–1418.
- [4] D. E. Quevedo, A. Ahlén, and K. H. Johansson, "State estimation over sensor networks with correlated wireless fading channels," *IEEE Trans. Autom. Control*, vol. 58, no. 3, pp. 581–593, Mar. 2013.
- [5] Y. Mo and B. Sinopoli, "Kalman filtering with intermittent observations: Tail distribution and critical value," *IEEE Trans. Autom. Control*, vol. 57, no. 3, pp. 677–689, Mar. 2012.
- [6] L. Shi, M. Epstein, and R. M. Murray, "Kalman filtering over a packet-dropping network: A probabilistic perspective," *IEEE Trans. Autom. Control*, vol. 55, no. 3, pp. 594–604, Mar. 2010.
- [7] L. Schenato, "Optimal estimation in networked control systems subject to random delay and packet drop," *IEEE Trans. Autom. Control*, vol. 53, no. 5, pp. 1311–1317, Jun. 2008.
- [8] M. Huang and S. Dey, "Stability of Kalman filtering with Markovian packet losses," *Automatica*, vol. 43, pp. 598–607, 2007.
- [9] D. E. Quevedo, A. Ahlén, A. S. Leong, and S. Dey, "On Kalman filtering over fading wireless channels with controlled transmission powers," *Automatica*, vol. 48, no. 7, pp. 1306–1316, Jul. 2012.
- [10] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inf. Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.
- [11] T. Ho and D. S. Lun, *Network Coding: An Introduction*. Cambridge, United Kingdom: Cambridge University Press, 2008.

- [12] D. E. Quevedo, A. Ahlén, and J. Østergaard, "Energy efficient state estimation with wireless sensors through the use of predictive power control and coding," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4811–4823, Sep. 2010.
- [13] D. E. Quevedo, J. Østergaard, and A. Ahlén, "Power control and coding formulation for state estimation with wireless sensors," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 2, pp. 413–427, Mar. 2014.
- [14] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
- [15] A. Host-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2020–2040, 2005.
- [16] T. S. Rappaport, R. W. Heath Jr., R. C. Daniels, and J. N. Murdock, *Millimeter Wave Wireless Communications*. New Jersey: Prentice Hall, 2014.
- [17] U. Kumar, V. Gupta, and J. N. Laneman, "Sufficient conditions for stabilizability over Gaussian relay and cascade channels," in *Proc. CDC*, Atlanta, GA, Dec. 2010.
- [18] T. A. Johansen, A. Zolich, T. Hansen, and A. J. Sørensen, "Unmanned aerial vehicle as communication relay for autonomous underwater vehicle - field tests," in *Proc. IEEE Globecom Workshop - Wireless Networking and Control for Unmanned Autonomous Vehicles*, Austin, TX, 2014.
- [19] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft, "XORs in the air: Practical wireless network coding," *IEEE/ACM Trans. Netw.*, vol. 16, no. 3, pp. 497–510, Jun. 2008.
- [20] D. Hui and D. L. Neuhoff, "Asymptotic analysis of optimal fixed-rate uniform scalar quantization," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 957–977, 2001.
- [21] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.
- [22] Y. Xu and J. P. Hespanha, "Estimation under uncontrolled and controlled communications in networked control systems," in *Proc. IEEE Conf. Decision and Control*, Seville, Spain, December 2005, pp. 842–847.
- [23] J. G. Proakis and M. Salehi, *Digital Communications*, 5th ed. New York: McGraw-Hill, 2008.
- [24] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1468–1489, Jul. 1999.
- [25] K. Gatsis, A. Ribeiro, and G. J. Pappas, "Optimal power management in wireless control systems," in *Proc. ACC*, Washington, DC, Jun. 2013, pp. 1565–1572.
- [26] HART Communication Foundation, "Control with WirelessHART," 2009.
- [27] W. J. Dally and R. C. Harting, *Digital Design: A Systems Approach*. Cambridge, UK: Cambridge University Press, 2012.
- [28] Z. Tang, H. Wang, Q. Hu, and L. Hai, "How network coding benefits converge-cast in wireless sensor netwrk," in *Proc. IEEE VTC Fall*, Quebec City, Canada, Sep. 2012, pp. 1–5.
- [29] G. Rajalingham, Q.-D. Ho, and T. Le-Ngoc, "Random linear network coding for converge-cast smart grid wireless networks," in *Proc. QBSC*, Kingston, ON, Jun. 2014, pp. 208–212.
- [30] M. Locatelli and F. Schoen, *Global Optimization: Theory, Algorithms and Applications*. Philadelphia: SIAM, 2013.
- [31] D. Dolz, D. E. Quevedo, I. Peñarrocha, and R. Sanchis, "Performance vs complexity trade-offs for Markovian networked jump estimators," in *Proc. IFAC World Congress*, Cape Town, South Africa, Aug. 2014, pp. 7412–7417.